

## VARIATIONAL REPRESENTATIONS OF VARADHAN FUNCTIONALS

HAROLD BELL AND WLODZIMIERZ BRYC

(Communicated by Claudia Neuhauser)

ABSTRACT. Motivated by the theory of large deviations, we introduce a class of non-negative non-linear functionals that have a variational “rate function” representation.

### 1. INTRODUCTION

Let  $(\mathbf{X}, d)$  be a Polish space with metric  $d(\cdot)$  and let  $\mathbf{C}_b(\mathbf{X})$  denote the space of all bounded continuous functions  $F : \mathbf{X} \rightarrow \mathbb{R}$ . In his work on large deviations of probability measures  $\mu_n$ , Varadhan [12] introduced a class of non-linear functionals  $\mathbb{L}$  defined by

$$(1) \quad \mathbb{L}(F) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \int_{\mathbf{X}} \exp(nF(\mathbf{x})) d\mu_n$$

and used the large deviations principle of  $\mu_n$  to prove the variational representation

$$(2) \quad \mathbb{L}(F) = L_0 + \sup_{\mathbf{x} \in \mathbf{X}} \{F(\mathbf{x}) - \mathbb{I}(\mathbf{x})\},$$

where  $\mathbb{I} : \mathbf{X} \rightarrow [0, \infty]$  is the *rate function* governing the large deviations, and  $L_0 := \mathbb{L}(0) = 0$ .

Several authors [1, 3, 4, 9, 10, 11] abstracted non-probabilistic components from the theory of large deviations. In particular, in [3] (see also [10, Theorem 3.1]) we give conditions which imply the rate function representation (2) when the limit (1) exists, and we show that the rate function is determined from the dual formula

$$(3) \quad \mathbb{I}(\mathbf{x}) = \mathbb{L}(0) + \sup_{F \in \mathbf{C}_b(\mathbf{X})} \{F(\mathbf{x}) - \mathbb{L}(F)\}.$$

In fact, one can reverse Varadhan’s approach, and show that large deviations of probability measures  $\mu_n$  follow from the variational representation (2) for (1) (see [8, Theorem 1.2.3]). In this context we have  $\mu_n(\mathbf{X}) = 1$  which implies  $\mathbb{L}(0) = 0$  in (3) and correspondingly  $L_0 = 0$  in (2).

“Asymptotic values” in [3] are essentially what we call Varadhan Functionals here; the theorems in that paper are not entirely satisfying because the assumptions are in terms of the underlying probability measures. In this paper we present a more

---

Received by the editors June 11, 1999 and, in revised form, November 10, 1999.

2000 *Mathematics Subject Classification*. Primary 60F10.

*Key words and phrases*. Large deviation, Čech-Stone compactification, Varadhan functionals, rate functions.

satisfying approach which relies on the theory of probability for motivation purposes only.

**Definition 1.1.** A function  $\mathbb{L} : \mathbf{C}_b(\mathbf{X}) \rightarrow \mathbb{R}$  is a Varadhan Functional if the following conditions are satisfied:

- (4) If  $F \leq G$ , then  $\mathbb{L}(F) \leq \mathbb{L}(G)$  for all  $F, G \in \mathbf{C}_b(\mathbf{X})$ ,  
 (5)  $\mathbb{L}(F + \text{const}) = \mathbb{L}(F) + \text{const}$  for all  $F \in \mathbf{C}_b(\mathbf{X})$ ,  $\text{const} \in \mathbb{R}$ .

Expression (1) provides an example of Varadhan Functional, if the limit exists.

Another example is given by variational representation (2).

Condition (4) is equivalent to  $\mathbb{L}(F \vee G) \geq \mathbb{L}(F) \vee \mathbb{L}(G)$ , where  $a \vee b$  denotes the maximum of two numbers. Varadhan Functionals like (1) satisfy a stronger condition.

**Definition 1.2.** A Varadhan Functional  $\mathbb{L}$  is *maximal* if  $\mathbb{L}(\cdot)$  is a lattice homomorphism

$$(6) \quad \mathbb{L}(F \vee G) = \mathbb{L}(F) \vee \mathbb{L}(G).$$

It is easy to see that each Varadhan Functional  $\mathbb{L}(\cdot)$  satisfies the Lipschitz condition  $|\mathbb{L}(F) - \mathbb{L}(G)| \leq \|F - G\|_\infty$ ; compare (9). Thus  $\mathbb{L}$  is a continuous mapping from the Banach space  $\mathbf{C}_b(\mathbf{X})$  of all bounded continuous functions into the real line. We will need the following stronger continuity assumption, motivated by the definition of the countable additivity of measures.

**Definition 1.3.** A Varadhan Functional is  *$\sigma$ -continuous* if the following condition is satisfied:

$$(7) \quad \text{If } F_n \searrow 0, \text{ then } \mathbb{L}(F_n) \rightarrow \mathbb{L}(0).$$

Notice that if  $\mathbf{X}$  is compact, then by Dini's theorem and the Lipschitz property, all Varadhan Functionals are  $\sigma$ -continuous.

Maximal Varadhan Functionals are convex; this follows from the proof of Theorem 2.1, which shows that formula (2) holds true for all Varadhan Functionals when the supremum is extended to all  $\mathbf{x}$  in the Čech-Stone compactification of  $\mathbf{X}$ .

A simple example of convex and maximal but not  $\sigma$ -continuous Varadhan Functional is  $\mathbb{L}(F) = \limsup_{x \rightarrow \infty} F(x)$ , where  $F \in \mathbf{C}_b(\mathbb{R})$ . This Varadhan Functional cannot be represented by variational formula (2). Indeed, (2) implies that  $\mathbb{L}(\mathbf{x}) \geq F(\mathbf{x}) - \mathbb{L}(F) = F(\mathbf{x})$  for all  $F \in \mathbf{C}_b(\mathbb{R})$  that vanish at  $\infty$ ; hence  $\mathbb{L}(\mathbf{x}) = \infty$  for all  $\mathbf{x} \in \mathbb{R}$  and (2) gives  $\mathbb{L}(F) = -\infty$  for all  $F \in \mathbf{C}_b(\mathbb{R})$ , a contradiction.

An example of a convex and  $\sigma$ -continuous but not maximal Varadhan Functional is  $\mathbb{L}(F) = \log \int_{\mathbf{X}} \exp F(\mathbf{x}) \nu(d\mathbf{x})$ , where  $\nu$  is a finite non-negative measure.

## 2. VARIATIONAL REPRESENTATIONS

The main result of this paper is the following.

**Theorem 2.1.** *If a maximal Varadhan Functional  $\mathbb{L} : \mathbf{C}_b(\mathbf{X}) \rightarrow \mathbb{R}$  is  $\sigma$ -continuous, then there is  $L_0 \in \mathbb{R}$  such that variational representation (2) holds true and the rate function  $\mathbb{I} : \mathbf{X} \rightarrow [0, \infty]$  is given by the dual formula (3). Furthermore,  $\mathbb{I}(\cdot)$  is a tight rate function: sets  $\mathbb{I}^{-1}([0, a]) \subset \mathbf{X}$  are compact for all  $a > 0$ .*

The next result is closely related to Bryc [3, Theorem T.1.1] and Deuschel & Stroock [6, Theorem 5.1.6]. Denote by  $\mathcal{P}(\mathbf{X})$  the metric space (with Prokhorov

metric) of all probability measures on a Polish space  $\mathbf{X}$  with the Borel  $\sigma$ -field generated by all open sets.

**Theorem 2.2.** *If a convex Varadhan Functional  $\mathbb{L} : \mathbf{C}_b(\mathbf{X}) \rightarrow \mathbb{R}$  is  $\sigma$ -continuous, then there is a lower semicontinuous function  $\mathbb{J} : \mathcal{P}(\mathbf{X}) \rightarrow [0, \infty]$  and a constant  $L_0$  such that*

$$(8) \quad \mathbb{L}(F) = L_0 + \sup_{\mu \in \mathcal{P}} \left\{ \int F d\mu - \mathbb{J}(\mu) \right\}$$

for all bounded continuous functions  $F$ .

A well-known example in large deviations is the convex  $\sigma$ -continuous functional  $\mathbb{L}(F) := \log \int \exp F(\mathbf{x}) \nu(d\mathbf{x})$  with the rate function in (8) given by the relative entropy functional

$$\mathbb{J}(\mu) = \begin{cases} \int \log \frac{d\mu}{d\nu} d\mu & \text{if } \mu \ll \nu \text{ is absolutely continuous,} \\ \infty & \text{otherwise.} \end{cases}$$

*Remark 2.1.* Deuschel & Stroock [6, Section 5.1] consider convex functionals  $\Phi : \mathbf{C}_b(\mathbf{X}) \rightarrow \mathbb{R}$  such that  $\Phi(\text{const}) = \text{const}$ . Such functionals satisfy condition (5). Indeed, write  $F + \text{const}$  as a convex combination

$$F + \text{const} = (1 - \theta)F + \frac{\theta}{2} \left( \frac{2\text{const}}{\theta} \right) + \frac{\theta}{2}(2F),$$

where  $0 < \theta < 1$ . Using convexity and  $\Phi(\text{const}) = \text{const}$  we get  $\Phi(F + \text{const}) \leq \Phi(F) + \text{const} + \theta \left( \frac{\Phi(2F)}{2} - \Phi(F) \right)$ . Since  $\theta > 0$  is arbitrary this proves that  $\Phi(F + \text{const}) \leq \Phi(F) + \text{const}$ . By routine symmetry considerations (replacing  $F \mapsto F - \text{const}$ , and then  $\text{const} \mapsto -\text{const}$ ), (5) follows.

### 3. PROOFS

Let  $L_0 := \mathbb{L}(0)$ . Passing to  $\mathbb{L}'(F) := \mathbb{L}(F) - L_0$  if necessary, without losing generality, we assume  $\mathbb{L}(0) = 0$ .

**Lemma 3.1.** *Let  $\hat{\mathbf{X}}$  be a compact Hausdorff space. Suppose  $\mathbf{X} \subset \hat{\mathbf{X}}$  is a separable metric space in the relative topology. If  $\mathbf{x}_0 \in \hat{\mathbf{X}} \setminus \mathbf{X}$ , then there are bounded continuous functions  $F_n : \hat{\mathbf{X}} \rightarrow \mathbb{R}$  such that*

- (i)  $F_n(\mathbf{x}) \searrow 0$  for all  $\mathbf{x} \in \mathbf{X}$ ,
- (ii)  $F_n(\mathbf{x}_0) = 1$  for all  $n \in \mathbb{N}$ .

*Proof.* Since  $\hat{\mathbf{X}}$  is Hausdorff, for every  $\mathbf{x} \in \mathbf{X}$  there is an open set  $U_{\mathbf{x}} \ni \mathbf{x}$  such that its closure  $\bar{U}_{\mathbf{x}}$  does not contain  $\mathbf{x}_0$ .

By Lindelöf property for separable metric space  $\mathbf{X}$ , there is a countable subcover  $\{U_n\}$  of  $\{U_{\mathbf{x}}\}$ .

A compact Hausdorff space  $\hat{\mathbf{X}}$  is normal. So there are continuous functions  $\phi_n : \hat{\mathbf{X}} \rightarrow \mathbb{R}$  such that  $\phi_n|_{\bar{U}_n} = 0$  and  $\phi_n(\mathbf{x}_0) = 1$ .

To end the proof take  $F_n(\mathbf{x}) = \min_{1 \leq k \leq n} \phi_k(\mathbf{x})$ . □

The following lemma is contained implicitly in [3, Theorem T.1.2].

**Lemma 3.2.** *Theorem 2.1 holds true for compact  $\mathbf{X}$ .*

*Proof.* Let  $\mathbb{I}(\cdot)$  be defined by (3). Thus  $\mathbb{I}(\mathbf{x}) \geq F(\mathbf{x}) - \mathbb{L}(F)$  which implies  $\mathbb{L}(F) \geq \sup_{\mathbf{x} \in \mathbf{X}} \{F(\mathbf{x}) - \mathbb{I}(\mathbf{x})\}$ . To end the proof we need therefore to establish the converse inequality. Fix a bounded continuous function  $F \in \mathbf{C}_b(\mathbf{X})$  and  $\epsilon > 0$ . Let  $s = \sup_{\mathbf{x} \in \mathbf{X}} \{F(\mathbf{x}) - \mathbb{I}(\mathbf{x})\}$ . Clearly  $F(\mathbf{x}) - \mathbb{I}(\mathbf{x}) \leq s \leq \mathbb{L}(F)$ . By (3) again, for every  $\mathbf{x} \in \mathbf{X}$ , there is  $F_{\mathbf{x}} \in \mathbf{C}_b(\mathbf{X})$  such that  $\mathbb{I}(\mathbf{x}) < F_{\mathbf{x}}(\mathbf{x}) - \mathbb{L}(F_{\mathbf{x}}) + \epsilon$ . Therefore

$$F(\mathbf{x}) \leq s + \mathbb{I}(\mathbf{x}) < s + \epsilon + F_{\mathbf{x}}(\mathbf{x}) - \mathbb{L}(F_{\mathbf{x}}).$$

This means that the sets  $U_{\mathbf{x}} = \{\mathbf{y} \in \mathbf{X} : F(\mathbf{y}) - F_{\mathbf{x}}(\mathbf{y}) < s + \epsilon - \mathbb{L}(F_{\mathbf{x}})\}$  form an open covering of  $\mathbf{X}$ . Using compactness of  $\mathbf{X}$ , we choose a finite covering  $U_{\mathbf{x}(1)}, \dots, U_{\mathbf{x}(k)}$ . Then, writing  $F_i = F_{\mathbf{x}(i)}$  we have

$$F(\mathbf{x}) < \max_{1 \leq i \leq k} \{F_i(\mathbf{x}) - \mathbb{L}(F_i)\} + s + \epsilon$$

for all  $\mathbf{x} \in \mathbf{X}$ .

Using (4), (5), and (6) we have

$$\begin{aligned} \mathbb{L}(F) &\leq \mathbb{L}\left(\max_{1 \leq i \leq k} \{F_i - \mathbb{L}(F_i)\} + s + \epsilon\right) \\ &= \mathbb{L}\left(\max_i \{F_i - \mathbb{L}(F_i)\}\right) + s + \epsilon \\ &= \max_i \{\mathbb{L}(F_i - \mathbb{L}(F_i))\} + s + \epsilon. \end{aligned}$$

Since (5) implies  $\mathbb{L}(F_i - \mathbb{L}(F_i)) = \mathbb{L}(F_i) - \mathbb{L}(F_i) = 0$  this shows that  $s \leq \mathbb{L}(F) < s + \epsilon$ . Therefore  $\mathbb{L}(F) = s$ , proving (2).  $\square$

*Proof of Theorem 2.1.* Let  $\hat{\mathbf{X}}$  be the Čech-Stone compactification of  $\mathbf{X}$ . Since the inclusion  $\mathbf{X} \subset \hat{\mathbf{X}}$  is continuous, we define  $\hat{\mathbb{L}} : \mathbf{C}_b(\hat{\mathbf{X}}) \rightarrow \mathbb{R}$  by  $\hat{\mathbb{L}}(\hat{F}) := \mathbb{L}(\hat{F}|_{\mathbf{X}})$ . It is clear that  $\hat{\mathbb{L}}$  is a maximal Varadhan Functional, so by Lemma 3.2 there is  $\mathbb{I} : \hat{\mathbf{X}} \rightarrow [0, \infty]$  such that  $\hat{\mathbb{L}}(\hat{F}) = \sup\{\hat{F}(\mathbf{x}) - \mathbb{I}(\mathbf{x}) : \mathbf{x} \in \hat{\mathbf{X}}\}$ .

Using  $\sigma$ -continuity (7) it is easy to check that  $\mathbb{I}(\mathbf{x}) = \infty$  for all  $\mathbf{x} \in \hat{\mathbf{X}} \setminus \mathbf{X}$ . Indeed, given  $\mathbf{x}_0 \in \hat{\mathbf{X}} \setminus \mathbf{X}$  by Lemma 3.1 there are  $F_n \in \mathbf{C}_b(\hat{\mathbf{X}})$  such that  $F_n \searrow 0$  on  $\mathbf{X}$ , but  $F_n(\mathbf{x}_0) = C > 0$ . Then from (3) we get  $\mathbb{I}(\mathbf{x}_0) \geq \hat{\mathbb{L}}(0) + F_n(\mathbf{x}_0) - \hat{\mathbb{L}}(F_n) \rightarrow \hat{\mathbb{L}}(0) + C$ . Since  $C > 0$  is arbitrary,  $\mathbb{I}(\mathbf{x}_0) = \infty$ .

This shows that  $\hat{\mathbb{L}}(\hat{F}) = \sup\{\hat{F}(\mathbf{x}) - \mathbb{I}(\mathbf{x}) : \mathbf{x} \in \mathbf{X}\}$  for all  $\hat{F} \in \mathbf{C}_b(\hat{\mathbf{X}})$ . It remains to observe that since  $\hat{\mathbf{X}}$  is a Čech-Stone compactification, every function  $F \in \mathbf{C}_b(\mathbf{X})$  is a restriction to  $\mathbf{X}$  of some  $\hat{F} \in \mathbf{C}_b(\hat{\mathbf{X}})$  (see [7, IV.6.22]). Therefore (2) holds true for all  $F \in \mathbf{C}_b(\mathbf{X})$ .

To prove that the rate function is tight, suppose that there is  $a > 0$  such that  $\mathbb{I}^{-1}[0, a]$  is not compact. Then there is  $\delta > 0$  and a sequence  $\mathbf{x}_n \in \mathbf{X}$  such that  $\mathbb{I}(\mathbf{x}_n) \leq a$ , and  $d(\mathbf{x}_m, \mathbf{x}_n) > \delta$  for all  $m \neq n$ . Since Polish spaces have Lindelöf property, there is a countable number of open balls of radius  $\delta/2$  which cover  $\mathbf{X}$ . For  $k = 1, 2, \dots$ , denote by  $B_k \ni \mathbf{x}_k$  one of the balls that contain  $\mathbf{x}_k$ , and let  $\phi_k$  be a bounded continuous function such that  $\phi_k(\mathbf{x}_k) = 2a$  and  $\phi_k = 0$  on the complement of  $B_k$ . Then  $F_n = \max_{k \geq n} \phi_k \searrow 0$  pointwise. On the other hand (2) implies  $\mathbb{L}(F_n) \geq L_0 + F_n(\mathbf{x}_n) - \mathbb{I}(\mathbf{x}_n) \geq L_0 + a$ , contradicting (7).  $\square$

**Lemma 3.3.** *If  $\mathbb{L}(\cdot)$  is a Varadhan Functional, then*

$$(9) \quad \inf_{\mathbf{x} \in \mathbf{X}} \{F(\mathbf{x}) - G(\mathbf{x})\} \leq \mathbb{L}(F) - \mathbb{L}(G).$$

*Proof.* Let  $\text{const} = \inf_{\mathbf{x}} \{F(\mathbf{x}) - G(\mathbf{x})\}$ . Clearly,  $F \geq G + \text{const}$ . By positivity condition (4) this implies  $\mathbb{L}(F) \geq \mathbb{L}(G + \text{const}) = \mathbb{L}(G) + \text{const}$ .  $\square$

The next lemma is implicitly contained in the proof of [3, Theorem T.1.1]. Let  $\mathcal{P}_a(\mathbf{X})$  denote all regular finitely-additive probability measures on  $\mathbf{X}$  with the Borel field.

**Lemma 3.4.** *If  $\mathbb{L}(\cdot)$  is a convex Varadhan Functional on  $\mathbf{C}_b(\mathbf{X})$ , then there exist a lower semicontinuous function  $\mathbb{J} : \mathcal{P}_a(\mathbf{X}) \rightarrow [0, \infty]$  such that*

$$(10) \quad \mathbb{L}(F) = \mathbb{L}(0) + \sup\{\mu(F) - \mathbb{J}(\mu) : \mu \in \mathcal{P}_a(\mathbf{X})\},$$

*and the supremum is attained.*

*Proof.* Let  $\mathbb{J}(\cdot)$  be defined by

$$(11) \quad \mathbb{J}(\mu) = \mathbb{L}(0) + \sup\{\mu(F) - \mathbb{L}(F) : F \in \mathbf{C}_b(\mathbf{X})\}$$

and fix  $F_0 \in \mathbf{C}_b(\mathbf{X})$ . Recall that throughout this proof we assume  $\mathbb{L}(0) = 0$ .

By the definition of  $\mathbb{J}(\cdot)$ , we need to show that

$$(12) \quad \mathbb{L}(F_0) = \sup_{\mu} \inf_F \{\mu(F_0) - \mu(F) + \mathbb{L}(F)\},$$

where the supremum is taken over all  $\mu \in \mathcal{P}_a(\mathbf{X})$  and the infimum is taken over all  $F \in \mathbf{C}_b(\mathbf{X})$ . Moreover, since (11) implies that  $\mathbb{J}(\mu) \geq \mu(F_0) - \mathbb{L}(F_0)$  for all  $\mu \in \mathcal{P}_a(\mathbf{X})$ , then  $\mathbb{L}(F_0) \geq \sup_{\mu} \inf_F \{\mu(F_0) - \mu(F) + \mathbb{L}(F)\}$ . Hence to prove (12), it remains to show that there is  $\nu \in \mathcal{P}_a(\mathbf{X})$  such that

$$(13) \quad \mathbb{L}(F_0) \leq \nu(F_0) - \nu(F) + \mathbb{L}(F) \text{ for all } F \in \mathbf{C}_b(\mathbf{X})$$

(also, for this  $\nu$ , the supremum in (10) will be attained). To find  $\nu$ , consider the following sets. Let

$$\mathcal{M} = \{F \in \mathbf{C}_b(\mathbf{X}) : \inf_{\mathbf{x}} [F(\mathbf{x}) - F_0(\mathbf{x})] > 0\}$$

and let  $\mathcal{N}$  be a set of all finite convex combinations of functions  $g(\mathbf{x})$  of the form  $g(\mathbf{x}) = F(\mathbf{x}) + \mathbb{L}(F_0) - \mathbb{L}(F)$ , where  $F \in \mathbf{C}_b(\mathbf{X})$ .

It is easily seen from the definitions that  $\mathcal{M}$  and  $\mathcal{N}$  are convex; also  $\mathcal{M} \subset \mathbf{C}_b(\mathbf{X})$  is non-empty since  $1 + F_0 \in \mathcal{M}$ , and open since  $\{F : \inf_{\mathbf{x}} [F(\mathbf{x}) - F_0(\mathbf{x})] \leq 0\} \subset \mathbf{C}_b(\mathbf{X})$  is closed. Furthermore,  $\mathcal{M}$  and  $\mathcal{N}$  are disjoint. Indeed, take arbitrary

$$\mathcal{N} \ni g = \sum \alpha_k F_k + \mathbb{L}(F_0) - \sum \alpha_k \mathbb{L}(F_k).$$

Then

$$\begin{aligned} & \inf_x \{g(\mathbf{x}) - F_0(\mathbf{x})\} \\ &= \inf_x \left\{ \sum \alpha_k F_k(\mathbf{x}) - F_0(\mathbf{x}) \right\} - \sum \alpha_k \mathbb{L}(F_k) + \mathbb{L}(F_0) \\ &\leq \inf_x \left\{ \sum \alpha_k F_k(\mathbf{x}) - F_0(\mathbf{x}) \right\} - \mathbb{L}\left(\sum \alpha_k F_k\right) + \mathbb{L}(F_0) \leq 0, \end{aligned}$$

where the first inequality follows from the convexity of  $\mathbb{L}(\cdot)$  and the second one follows from (9) applied to  $F = \sum \alpha_k F_k(\mathbf{x})$  and  $G = F_0$ .

Therefore  $\mathcal{M}$  and  $\mathcal{N}$  can be separated, i.e. there is a non-zero linear functional  $f^* \in \mathbf{C}_b^*(\mathbf{X})$  such that for some  $\alpha \in \mathbb{R}$

$$(14) \quad f^*(\mathcal{N}) \leq \alpha < f^*(\mathcal{M})$$

(see e.g. [7, V. 2. 8]).

**Claim:**  $f^*$  is non-negative.

Indeed, it is easily seen that  $F_0(\cdot)$  belongs to  $\mathcal{N}$ , and, as a limit of  $\epsilon + F_0(\mathbf{x})$  as  $\epsilon \rightarrow 0$ ,  $F_0$  is also in the closure of  $\mathcal{M}$ . Therefore by (14) we have  $\alpha = f^*(F_0)$ . To end the proof take arbitrary  $F$  with  $\inf_{\mathbf{x}} F(\mathbf{x}) > 0$ . Then  $F + F_0 \in \mathcal{M}$  and by (14)

$$f^*(F) = f^*(F + F_0) - f^*(F_0) > \alpha - f^*(F_0) = 0.$$

This ends the proof of the claim.

Without losing generality, we may assume  $f^*(1) = 1$ ; then it is well known (see e.g. [2, Ch. 2 Section 4 Theorem 1]) that  $f^*(F) = \nu(F)$  for some  $\nu \in \mathcal{P}_a(\mathbf{X})$ ; for regularity of  $\nu$  consult [7, IV.6.2]. It remains to check that  $\nu$  satisfies (13). To this end observe that since  $F + \mathbb{L}(F_0) - \mathbb{L}(F) \in \mathcal{N}$ , by (14) we have  $\nu(F) + \mathbb{L}(F_0) - \mathbb{L}(F) \leq \alpha = \nu(F_0)$  for all  $F \in \mathbf{C}_b(\mathbf{X})$ . This ends the proof of (10).  $\square$

*Proof of Theorem 2.2.* Lemma 3.4 gives the variational representation (10) with the supremum taken over a too large set. To end the proof we will show that  $\mathbb{J}(\mu) = \infty$  on measures  $\mu$  that fail to be countably-additive.

Suppose that  $\mu$  is additive but not countably additive. Then Daniell-Stone theorem implies that there is  $\delta > 0$  and a sequence  $F_n \searrow 0$  of bounded continuous functions such that  $\int F_n d\mu > \delta > 0$  for all  $n$ . By (11) and  $\sigma$ -continuity,  $\mathbb{J}(\mu) \geq \mathbb{L}(0) + C \int F_n d\mu - \mathbb{L}(CF_n) \geq \mathbb{L}(0) + C\delta - \mathbb{L}(CF_n) \rightarrow \mathbb{L}(0) + C\delta$ . Since  $C > 0$  is arbitrary, then  $\mathbb{J}(\mu) = \infty$  for all  $\mu$  that are additive but not countably-additive. Thus (10) implies (8).  $\square$

## REFERENCES

1. M. Akian, Densities of idempotent measures and large deviations. Trans. Amer. Math. Soc. 351 (1999), 4515–4543. CMP 99:17
2. H. Bergström, *Weak convergence of measures*. Acad. Press, New York, 1982. MR **84m**:60027
3. W. Bryc, On the large deviation principle by the asymptotic value method. In: *Diffusion Processes and Related Problems in Analysis*, Vol. I, ed. M. Pinsky, Birkhäuser, Boston, 1990, 447–472.
4. A. de Acosta, Upper bounds for Large Deviations of Dependent Random Vectors. Zeitsch. Wahrscheinlichk. Theor. Verw. Gebiete 69 (1985), 551–565. MR **87f**:60036
5. A. Dembo & O. Zeitouni, *Large Deviations Techniques and Applications*. Jones and Bartlett, Boston, 1993. MR **95a**:60034
6. J-D. Deuschel & D. W. Stroock, *Large Deviations*. Pure and Applied Math vol. 137, Academic Press, Boston, 1989. MR **90h**:60026
7. N. Dunford & J. T. Schwartz, *Linear Operators I*. Interscience, New York, 1958. MR **90g**:47001a
8. P. Dupuis & R. S. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley, New York, 1997. MR **99f**:60057
9. G. L. O'Brien, Sequences of capacities, with connections to large-deviation theory. J. Theoret. Probab. 9 (1996), 19–35. MR **97f**:60065
10. G. O'Brien & W. Vervaat, Compactness in the theory of large deviations. Stoch. Processes Appl. 57 (1995), 1–10. MR **96a**:60030

11. A. Puhalskii, Large deviations of Semimartingales: a Maxingale Problem Approach I. Stochastics 61 (1997), 141–243. MR **98h**:60033
12. S. R. S. Varadhan, Asymptotic probabilities and differential equations. Comm. Pure Appl. Math. 19 (1966), 261–286. MR **34**:3083

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CINCINNATI, P.O. BOX 210025, CINCINNATI, OHIO 45221-0025

*E-mail address:* `bellh@math.uc.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CINCINNATI, P.O. BOX 210025, CINCINNATI, OHIO 45221-0025

*E-mail address:* `brycwz@email.uc.edu`